

## REMARKS

The above amendments to the above-captioned application along with the following remarks are being submitted as a response to the Official Action dated November 29, 2006. In view of the above amendments and the following remarks, the Examiner is respectfully requested to give due reconsideration to this application, to indicate the allowability of the claims, and to pass this case to issue.

### Status of the Claims

Claims 1, 3-4 and 6-8 are under consideration in this application. Claims 1, 4, 6 and 8 are being amended, as set forth in the above marked-up presentation of the claim amendments, in order to more particularly define and distinctly claim applicants' invention.

The claims are being amended to correct formal errors and/or to better recite or describe the features of the present invention as claimed. All the amendments to the claims are supported by the specification. Applicants hereby submit that no new matter is being introduced into the application through the submission of this response.

### Formality Rejection

The specification is objected to for reciting hyperlinks, and claims 6 and 8 were objected to for reciting "A program" which should be changed into "The program." As the specification and the claims are being amended as suggested by the Examiner, the withdrawal of the outstanding informality rejection is in order, and is therefore respectfully solicited.

### Prior Art Rejections

Claims 1, 3, 4 and 6-8 were rejected under 35 U.S.C. § 103 as being unpatentable over US Pat. App. Pub. No. 2002/0168664 issued into US. Pat. No. 6,876,930 to Murray et al. (hereinafter "Murray") in view of US Patent No. 6,519,592 to Getchius (hereinafter "Getchius"). This rejection has been carefully considered, but is most respectfully traversed in view of the newly submitted claims, as more fully discussed below.

The method of calculating the frequency of appearance of a keyword of the invention (for example, the embodiment depicted in Fig. 8), as now recited in claim 1, uses a first database in which information about a base sequence or an amino acid sequence is stored and a second database in which document data is stored (for example, per record or per document; Figs. 2-3 show the structure of first and second documents, and "*each file of text data 106 contains identifier 202 for identifying document data*" p. 5, last paragraph). The

method comprises: a first document extraction step (Steps 803+) for extracting a first document 106 (for example, “*a thesis describing the result of research into a particular base sequence*” p. 4, last paragraph) from said first database 105 (Fig. 2) which contains a base sequence or an amino acid sequence of a gene or protein of interest (e.g., “AGCT” p. 5, 2<sup>nd</sup> to last paragraph, 201 in Fig. 2, or 401 in Fig. 4) inputted by a user; an identifier extraction step (Steps 806+) for extracting an identifier 202 (e.g., PMID: P00005, Fig. 1; p. 4, last line; PubMed Identifier is a unique number assigned to each PubMed citation of life sciences and biomedical scientific journal articles) identifying document data (e.g., Fig. 2: “As seen in the example ....”) in said first document 106 from said extracted first document 106 which contains the base sequence or the amino acid sequence 201; a second document extraction step (Step 807) for extracting a second document 108 (for example, “*the data describes the result of molecular-biological study into a gene or protein*” p. 5, 1<sup>st</sup> paragraph) from said second database 107 (Fig. 3) which contains said extracted identifier 301 (same as 202, e.g., PMID: P00005); an appearance frequency calculation step (Step 810) for sequentially reading predefined keywords 503 (e.g., “axon midline choice point recognition” etc. listed in the last column of the table in Fig. 5) from a keyword table 110 (Fig. 5; “*The calculating unit 102 obtains keywords contained in the category table 110 in the third file system 109*” p. 6, lines 3-4) containing ***keywords of known functions or characteristics of genes or proteins*** from said first database, and for calculating a frequency of appearance (e.g., 2-1249 in Fig. 6) of each of said keywords 501-503 automatically and mechanically counting per document a number of extracted second documents containing said keywords 503 therein (“*Specifically, the number of files of extracted text data 108 in which each keyword appears or is used is calculated.*” p. 6, lines 5-6); and a displaying step for displaying a frequency of appearance of each of said keywords 501-503 in a corresponding position in said keyword table 110 (e.g., Figs. 6-7) thereby showing numbers of documents including said keywords to determine whether to select or change a research course (p. 3, lines 8-16). The keyword table 110 has a tree structure in which keywords 503 are stored such that they are classified according to categories 501, 502. The appearance frequency calculation step comprises a step for generating a frequency calculation result table (e.g., Fig. 7) of a tree structure, said table containing the frequency of appearance of a keyword 503 and the frequency of appearance of an upper-level category 501 to which the keyword belongs (“*In the illustrated example, the frequency of appearance of all of the keywords belonging to the category “cell recognition” is 196. This indicates that keywords belonging to the category “cell recognition” appear at* 6

*least once in 196 files of the text data contained in the second file system 107.” p. 6, last paragraph).*

The invention recited in claim 4 is directed to a program embedded in a storage medium for causing a computer to carry out the keyword frequency calculation method of claim 1.

*“The user can thus learn the frequency of each keyword related to the sequence AGCT in the text data 108 in the second file system 107 (p. 6, 2<sup>nd</sup> paragraph),” so as to find out the functions or characteristics of the particular gene or protein (p. 1, lines 18-19; P. 3, 2<sup>nd</sup> paragraph).*

In accordance with the invention, when a searcher wishes to know the functions or characteristics of a gene or protein with a particular sequence, the searcher can be provided with a list of keywords indicating the functions or characteristics of genes or proteins containing the particular sequence by entering the particular sequence information itself as a search key, the list showing the keywords in terms of the importance, or the frequency of appearance in document data (p. 3, 2<sup>nd</sup> paragraph). By viewing the keyword frequency table in Fig. 7, the researcher can know how much literature exists on the subject to know whether the subject has already been researched or studied enough so as to determine whether to select or change the research’s course (p. 3, lines 8-16).

Applicants respectfully contend that none of the cited references teaches or suggests such “a displaying step for displaying *a frequency of appearance* (e.g., 2-1249 in Fig. 6) of each of said keywords 501-503 in a corresponding position in said keyword table 110 (e.g., Figs. 6-7) thereby showing numbers of documents including said keywords to determine whether to select or change a research course” according to the invention.

In contrast, Murray just uses “journal titles” as keywords to score the times each of the journal titles was “*referenced in other journals*” to *rank* the *relative* impact of a give paper ([0101] Murray ‘664) as from 10-100 as in Table 1 (p. 10 of Murray ‘664), rather than “displaying *an exact frequency of appearance* (e.g., 2-1249 in Fig. 6) of each of said keywords 501-503 in a corresponding position in said keyword table 110” as does the present invention. Murray neither concerns “showing numbers of documents including said keywords to determine whether to select or change a research course” as does the present invention.

In addition, Murray’s “*journal titles*” are essentially different from the “*keywords of known functions or characteristics* of genes or proteins, such as “axon midline choice point recognition” etc. of the present invention.

Moreover, Murray extracts relationships between the journal articles so as to extracts relationships between genes (“*to provide further information about a particular DNA sequence to facilitate its identification as a candidate gene*” Abstract). On the other hand, the invention merely extracts relationships between the documents by consisting of limited steps.

Contrary to the Examiner’s assertion (p. 6, lines 16-18 of the outstanding Office Action) that the invention recited in claim 1 only uses the identifier (e.g., PMID) to identify document data, but not as the criteria to extract the relationship between genes. The invention recited claim 1 does not intend to have a user to view the second database to verify gene expression results as asserted by the Examiner. It appears that the Examiner refers to the second database of the invention as Murray’s “expert database (col. 26, lines 20-26)”. However, Murray’s “expert database” stores a list of interactions and a directed graph, rather than any document data as the second database of the present invention.

Getchius was relied upon by the Examiner (p. 6, 4th paragraph of the outstanding Office Action) to teach using a tree structure for the keyword table as recited in claims 2 and 5 of the invention which are now incorporated in to claims 1 and 4. However, Getchius is applied to a business listing (Abstract; e.g., ““*restaurant*” may be stored in a tree that includes the sub-category of “*ethnic restaurant*,” which may further include the sub-category “*Greek restaurant*. ”” col. 33, lines 19-21), rather than any “known functions or characteristics of genes or proteins” as the keywords in the keyword table 110 of the invention.

Even if, arguendo, one skilled in the art were motivated to combine the teachings in Murray and Getchius as suggested by the Examiner, such combined teachings would still fall short in fully meeting the Applicants’ claimed invention as set forth in claims 1 and 4, since neither reference teaches “a displaying step for displaying *a frequency of appearance* (e.g., 2-1249 in Fig. 6) of each of said keywords 501-503 in a corresponding position in said keyword table 110 (e.g., Figs. 6-7) thereby showing numbers of documents including said keywords to determine whether to select or change a research course” as in the invention. At most, Getchius shows numbers of restaurants of different cuisines to determine where to eat, rather than “to determine whether to select or change a research course” as the present invention.

Applicants contend that neither Murray, Getchius, nor their combinations teaches or discloses each and every feature of the present invention as recited in independent claims 1 and 4. As such, the present invention as now claimed is distinguishable and thereby allowable over the rejections raised in the Office Action. The withdrawal of the outstanding prior art rejections is in order, and is respectfully solicited.

Conclusion

In view of all the above, clear and distinct differences as discussed exist between the present invention as now claimed and the prior art reference upon which the rejections in the Office Action rely, Applicants respectfully contend that the prior art references cannot anticipate the present invention or render the present invention obvious. Rather, the present invention as a whole is distinguishable, and thereby allowable over the prior art.

Favorable reconsideration of this application is respectfully solicited. Should there be any outstanding issues requiring discussion that would further the prosecution and allowance of the above-captioned application, the Examiner is invited to contact the Applicants' undersigned representative at the address and telephone number indicated below.

Respectfully submitted,

Stanley P. Fisher  
Registration Number 24,344

Juan Carlos A. Marquez  
Registration Number 34,072

**REED SMITH LLP**  
3110 Fairview Park Drive, Suite 1400  
Falls Church, Virginia 22042  
(703) 641-4200

**February 28, 2007**

SPF/JCM/JT